Measuring Internet retail service quality using E-S-QUAL

Mohammed Rafiq, *Loughborough University, UK*
Xiaoming Lu, *Northumbria University, UK*
Heather Fulford, *The Robert Gordon University, UK*

Abstract Despite its acknowledged importance, there are few rigorous empirical studies examining Internet retail service quality. An exception is the development of the E-S-QUAL scale by Parasuraman, Zeithaml, and Malhotra (2005). Whilst E-S-QUAL demonstrated good psychometric properties in the original study, the scale lacks external validation. This paper presents a reassessment and validation of the E-S-QUAL in the context of the Internet grocery sector. Data were collected via a web-based cross-sectional survey using self-administered questionnaires distributed to online grocery shoppers. A total of 491 usable questionnaires were received. The results show that there are potential discriminant validity problems with the Efficiency and System Availability dimensions of E-S-QUAL. Further analysis shows that a second-order, three-factor model of E-S-QUAL, consisting of Efficiency, System Availability, and Fulfilment, provides the best fit to the data in this study. Privacy is shown to be the least important dimension for the data set in this study.

Keywords Internet retail service quality; e-service quality; E-S-QUAL

Introduction and background

With the increasing importance of Internet retailing, service quality in the online environment has been recognised as an important factor in determining the success or failure of e-commerce ventures (Santos, 2003; Yang, 2001; Zeithaml, Parasuraman, & Malhotra, 2002). A number of existing studies on e-service quality have attempted to identify the elements that define customers' perception of service quality, and to build models that outline the differences between customers' expectations and the real service experience (Janda, Trocchia, & Gwinner, 2002; Zeithaml et al., 2002). Much of the early empirical research on Internet retail service quality, the focus of this study, concentrated on developing measures for the evaluation of websites. However, Wolfinbarger and Gilly (2003) argue that measuring e-service quality should go beyond the website interface. This is because a customer's online buying experience consists of everything from information search, product

evaluation, decision making, the transaction, delivery, returns, and customer service. It is apparent that measures for evaluating just websites may not be sufficient for measuring service quality across various stages of the online retail service delivery. This is also in line with Parasuraman, Zeithaml, and Malhotra's (2005; henceforth, PZM) view, who state that the purpose of developing scales for e-service quality is to measure the whole experience of customers regarding the service received rather than to generate information for website designers.

There have been relatively few rigorous empirical studies examining Internet retail service quality to date. Examples of such studies include Barnes and Vidgen (2002); Janda et al. (2002); Loiacono, Watson, and Goodhue (2002); Wolfinbarger and Gilly (2003); and Yoo & Donthu (2001). However, many of these studies do not include all aspects of service quality (see, e.g., Boshoff, 2006). Collier and Bienstock's (2006) study, however, is an exception to this criticism. Their study proposes and tests an e-service quality conceptualisation that includes process, outcome, and recovery dimensions. However, as Fassnacht and Koese (2006) point out, what customers are looking for in the first instance is high quality service and not recovery. Good recovery may be required in some instances, but it is not the primary focus of what customers want. Therefore, it is better to treat service recovery as a separate dimension. This is, in fact, what PZM do by proposing a separate scale (E-RecS-Qual) dealing with service recovery issues.

One study that stands out amongst these is PZM's (2005) E-S-QUAL scale with its four dimensions of electronic service quality. Based on their preliminary work (Zeithaml, Parasuraman, & Malhotra, 2000) arguing that in any assessment of Internet service quality, the focus ought to be on all cues and encounters that occur before, during, and after the transaction, PZM developed a 22-item scale with four dimensions, namely: Efficiency (ease and speed of accessing and using the site, eight items); System Availability (correct technical functioning of the site, four items); Fulfilment (extent to which the site's promises about order delivery and item availability are fulfilled, seven items); and Privacy (degree to which the site is safe and protects customer information, three items). The E-S-QUAL measure stands out because it is rigorously conceptualised and systematically tested. PZM meticulously followed procedures for developing new scales, and the scale demonstrates good psychometric properties. The E-S-QUAL measure, however, lacks external validation. To date, we are aware of only one attempt at external validation of E-S-QUAL – by Boshoff (2006) in the Internet retailing context. Boshoff concludes that E-S-QUAL captures the essence of e-service quality. However, Boshoff calls for further research into the dimensionality of the construct. This is because, in Boshoff's study, a six-factor scale provided a better fit to the data than the original four-factor E-S-QUAL scale. The two extra dimensions result from splitting the Efficiency dimension into Efficiency and 'Website Speed', and the Fulfilment dimension into 'Delivery' and 'Reliability'. However, the respondents in Boshoff's study were the customers of a single firm, limiting the potential generalisability of the results.

Given the questions raised by the Boshoff's study over the dimensionality of E-S-QUAL and the importance of reliable and valid research measures, a reassessment and revalidation of the E-S-QUAL scale in the context of the UK Internet grocery sector is presented, employing a cross-sectional survey design rather than focusing on a particular focal retail organisation. Due to the importance of service quality in the success of e-retailers, external validation of e-service quality measures through replication is extremely important, particularly in cases where measures developed

in one country are intended for use in other countries. Replications not only help to determine the reliability and validity of newly developed measurement instruments but also help to define the scope and limits to their generalisability to other contexts (Hubbard, Vetter, & Little, 1998). In the case of Internet retailing, due to the intrinsic borderless nature of the Internet, it may be easily assumed that e-service quality measures are equally applicable internationally, when, in fact, they are not. Therefore, the validity of E-S-QUAL needs to be tested and established in a cross-national context in order to identify limitations that it may have with respect to its generalisability. The UK context of this study helps to assess the robustness of the E-S-QUAL scale in an international context and hence its generalisability beyond the original US context. The UK online grocery market is one of the most competitive in the world with three out of the four major chains (Tesco, Sainsbury's, and ASDA) as well as Waitrose and the pure Internet retailer, Ocado, operating in the area. The US online grocery market, on the other hand, is more of a niche market with a relatively small number of regional operators providing delivery mostly in urban areas.

Online grocery shopping is an interesting area for testing E-S-QUAL because shopping for groceries is largely a replenishment, low-involvement activity that is repeated at regular time intervals. Therefore online service quality is likely to be even more important because of the frequency of the transactions and the amount that customers spend on groceries is relatively high. Also, PZM included Wal-Mart shoppers as respondents in the development of E-S-QUAL, and this therefore provides a point of comparison for this study. The study also tests for the second-order formulation of the E-S-QUAL construct. The existence of a second-order model structure would provide a more parsimonious view of how customers perceive e-service quality. The study also tests the nomological validity of E-S-QUAL in this context and extends the nomological net by treating it as an antecedent to customer satisfaction and loyalty.

Research method

As this paper focuses on validating E-S-QUAL, PZM's (2005) 22-item four-dimensional E-S-QUAL scale was adapted in this study (see Table 2). However, instead of a five-point Likert scale used in the PZM study, a seven-point Likert scale was employed in this study to extend the range and variability of responses. To measure customer satisfaction in the online environment, the measure from Jones and Suh's (2000) study was adopted. In their study, online satisfaction, or e-satisfaction, is measured using three semantic differential items commonly used to measure customer satisfaction in offline contexts: the degree to which the consumer is satisfied/dissatisfied (e.g. Oliver, 1980; Zeithaml, Berry, & Parasuraman, 1996), feels pleased/displeased (e.g. Spreng & Mackoy, 1996) and is favourable/unfavourable towards the Internet grocery retailer. Loyalty was measured using Zeithaml et al.'s (1996) five-item scale. Both these scales also employed seven-point Likert scales for the reasons mentioned above.

The data presented in this paper were collected via a web-based survey using self-administered questionnaires. The questionnaires were distributed to online grocery shoppers using an Internet panel administered by a market research company after a pilot study of 100 respondents. A total of 519 responses were received within a week and 491 questionnaires remained for further analysis after data screening.

Table 1 Comparative descriptives of UK study and PZM's Wal-Mart sample.

Variable	UK study (%)	Wal-Mart USA (%)
Age in years		
<25	9.2	10
25–40	38.2	43
41–55	39.0	34
>55	13.5	13
Sex		
Male	49.3	22
Female	50.7	78
Highest level of education		
Secondary school/*high school*[1]	24	21
College	40	41
University/*college graduate, graduate school*[1]	36	37
Annual household income		
<£20,000/*<$25,000*[2]	25.2	15
£20,000–£29,999/*$25,000–$49,999*[2]	26.2	35
£30,000–£39,999/*$50,000–$74,999*[2]	19.8	32
≥£40,000/*$75,000 or more*[2]	28.8	17
Length of website use		
<3 months	8.9	19.0
3–6 months	8.2	20.0
6–12 months	14.0	29
≥12 months	68.9	32
Frequency of use		
4 times or less a month	98.4	83
>4 times a month	1.6	17

[1]Italics denote equivalent terminology used in the PZM study. [2]Italics denote ranges used in the PZM study.

Overall, the two samples are reasonably comparable (see Table 1). However, there are three main differences between the two studies in terms of the profile of the respondents. First, the sample of respondents in this study is more balanced in terms of the gender of the respondents (49% male, 51% female) compared with PZM (22% male, 78% female). Second, in terms of the length of patronage, the respondents who had shopped with their current internet retailer for 12 months or more was 69% in this study compared with 32% for Wal-Mart sample in the PZM study. PZM attributed the low figure for Wal-Mart to the fact that Wal-Mart's website had not been in existence for long at the time of their study. Third, PZM's study had a higher proportion of respondents (17%) who shopped more than four times a month compared with this study (1.6%).

In the sample, 69.7% shopped online with Tesco, 14.5% with ASDA, 10.6% with Sainsbury's, .6% with Waitrose, and 4.7% with a number of smaller operators such as Ocado and Foodferry. On average, 64% of the respondents had shopped with the offline store before trying the online store. That is, around 64% of the respondents were transferring their loyalty from their offline store to the online store. The most

loyal were Tesco shoppers, of whom 93% had shopped with the offline store before shopping with the online store. Because of the unequal numbers of respondents shopping with different e-retailers, the Levene test statistics were calculated to ensure that homogeneity of variance among groups had been achieved. All Levene test statistics were insignificant ($p > .10$), indicating homogeneity.

Method of analysis

As E-S-QUAL is an existing, theoretically supported scale, confirmatory factor analysis (CFA) was used to assess its unidimensionality, reliability, and validity. Whilst exploratory factor analysis (EFA) is often used before conducting CFA, Gerbing and Anderson (1988, p. 189) argue that because factors obtained via EFA are defined as the 'weighted sum of *all* observed variables', they do not represent the theoretical constructs underlying each set of indicators. Following the CFA, the nomological validity of E-S-QUAL was then assessed by testing its relationships with customer satisfaction and loyalty in a nomological net (Steenkamp & van Trijp, 1991).

Results

As recommended by Garver and Mentzer (1999), the analysis began by looking at each of the dimensions separately to assess whether each of the items loaded on the dimension that it was supposed to, and to assess the unidimensionality of the constructs. All the items loaded quite well on the appropriate dimensions. The standardised loadings ranged from .80 to .89 for the Efficiency dimension, .80 to .88 for System Availability, .64 to .86 for Fulfilment, and .75 to .83 for Privacy (see Table 2). However, two items of the Fulfilment dimensions had correlated errors; item FUL1: it 'delivers orders when promised'; and item FUL7: it 'makes accurate promises about delivery of products'. On reflection, these items do appear to be very close and, therefore, item FUL7 was arbitrarily deleted, as both items had the same standardised loadings. The high standardised loadings suggested that the constructs were unidimensional. The Cronbach's alphas ranged from .83 for the Privacy dimension to .96 for the Efficiency dimension. As all the coefficient alphas exceeded the conventionally recommended minimum value of .7 (Nunnally & Bernstein, 1994), suggesting that the measures are reliable. Given this, the analysis proceeded to the next stage – conducting CFA by looking at all the dimensions together. The standardised loadings were all above .7 except for item 5, Fulfilment, which had a loading of .64. Overall, the standardised loadings show a similar pattern to those obtained by PZM. The goodness-of-fit (GFI) statistics are also similarly good. In this study, however, RMSEA is significantly lower at .08 compared with .11 in PZM's data.

The inter-factor correlations ranged from .54 to .88 (see Table 3). In the PZM study, the inter-factor correlations ranged from .62 to .77. The high standardised loadings together with the high coefficient alphas provide support for E-S-QUAL's convergent validity. To assess discriminant validity, the same procedure as employed by PZM was followed, and each inter-factor correlation was fixed, one at time, and the CFA was re-estimated to examine the difference in the chi-square statistic for the original and the constrained models. Each of the CFAs produced a significant chi-square statistic ($\Delta\chi^2$ with 1 df), except for when the correlation between Efficiency

Table 2 First-order CFA for E-S-QUAL in the PZM and UK studies.

Factors	Item wording	PZM (2005) Wal-Mart USA loadings[1,2]	Cronbach's alpha for Wal-Mart	UK study loadings[2]	Cronbach's alpha for UK study
Efficiency	*The website of my Internet grocery store. . .*		.94		.96
EFF1[3]	. . .makes it easy to find what I need	.87		.87	
EFF2	. . .makes it easy to get anywhere on the site	.86		.87	
EFF3	. . .enables me to complete a transaction quickly	.81		.80	
EFF4	. . .has well-organised information	.88		.87	
EFF5	. . .loads its pages fast	.77		.81	
EFF6	. . .is simple to use	.81		.88	
EFF7	. . .enables me to get on to it quickly	.78		.84	
EFF8	. . .is well organised	.82		.89	
Systems availability			0.84		0.89
SYS1	. . .is always available for business	.78		.80	
SYS2	. . .launches and runs right away	.74		.88	
SYS3	. . .does not crash	.75		.81	
SYS4	. . .has pages that do not freeze after I enter my order information	.79		.80	
Fulfilment	*My Internet grocery store*		.94		.89
FUL1	. . .delivers orders when promised	.94		.78	
FUL2	. . .delivers items within a suitable time frame	.91		.86	
FUL3	. . .quickly delivers what I order	.85		.83	
FUL4	. . .sends out the items ordered	.88		.72	
FUL5	. . .has in stock the items the company claims to have	.74		.64	

(*Continued*)

Table 2 (Continued).

Factors	Item wording	PZM (2005) Wal-Mart USA loadings[1,2]	Cronbach's alpha for Wal-Mart	UK study loadings[2]	Cronbach's alpha for UK study
FUL6	. . .is truthful about its offering	.68		.78	
FUL7	. . .makes accurate promises about delivery times	.84		Item deleted	
Privacy			.83		.83
PRI1	. . .protects information about my web-shopping behaviour	.77		.79	
PRI2	. . .does not share my personal information with other sites	.83		.75	
PRI3	. . .protects information about my credit card	.78		.83	
GFI statistics					
Chi-square		739.86		755.31	
df		203		183	
CFI		.97		.93	
NFI		.96		.91	
RFI		.95		.90	
TLI		.96		.92	
RMSEA		.11		.08	

[1]The figures in this column are taken from table 5 (p. 225) of the PZM (2005) paper. [2]All loadings are significant at $p < .01$ in both studies. [3]All the items are aligned with items in the PZM paper.

Table 3 Inter-factor correlations between E-S-QUAL dimensions.

	Efficiency	System availability	Fulfilment	Privacy
Efficiency	1.00			
System	.88	1.00		
Availability	(.77)			
Fulfilment	.63	.59	1.00	
	(.68)	(.68)		
Privacy	.59	.54	.59	.54
	(.62)	(.64)	(.62)	(.64)

The inter-factor correlations for Wal-Mart in the PZM study are shown in brackets.

and System Availability was constrained to 1. This resulted in a chi-square difference value of 1.6, which is not statistically significant for 1 df, suggesting that there may be a problem with discriminant validity for the Efficiency and System Availability measures, as there is a high correlation of .88 between the two constructs (see Table 3). This finding contrasts with that of PZM who found that discriminant validity was supported for all four of their factors. However, the high correlation can arise if the constructs involved are theoretically related to a higher-order construct such as e-service quality (Bagozzi & Heatherton, 1994; Garver & Menzer, 1999). This is because the subcomponents of higher-order models contain a significant amount of shared variance resulting from their common relationship with the higher-order construct (Bagozzi & Heatherton, 1994). The next section examines whether e-service quality is indeed a higher-order construct as implied above and postulated by PZM (see PZM, fn 1, p. 220).

E-S-QUAL as a higher-order construct

There is plenty of support, both conceptual and empirical, for service quality as a higher-order construct, that is, a construct that has a number of sub-dimensions. It is usually conceptualised as a reflective second-order construct (Grönroos, 1984; Parasuraman, Zeithaml, & Berry, 1988; Rust & Oliver, 1994), but more recently third-order formulations have also been proposed and tested (Brady & Cronin, 2001; Dabholkar, Thorpe, & Rentz, 1996; Dagger, Sweeney, & Johnson, 2007). In the developing literature on e-service quality, a similar approach is emerging. For instance, Cristobal, Flavián, and Guinalíu (2007) provide evidence for a second-order construct for their perceived e-service quality construct. Similarly, Collier and Bienstock (2006) propose and test a second-order formulation of e-service quality. Fassnacht and Koese (2006), on the other hand, propose a third-order formulation of their QES (Quality of Electronic Services) scale with environment, delivery, and outcome quality as second-order dimensions. Although the results are not reported in detail, PZM did run a second-order CFA, modeling the first-order E-S-QUAL dimensions as reflective indicators of overall e-service quality. PZM report that for the second-order model, the factor loadings and fit statistics were similar to the first-order model. Boshoff (2006) did not test for a second-order formulation of E-S-QUAL, although the low inter-factor correlations (ranging from .35 to .68) suggest that a first-order model fits their data better.

However, in a footnote, PZM suggest that according to the criteria postulated by Jarvis, Mackenzie, and Podsakoff (2003), it may be better to treat the first-order dimensions as formative indicators of the second-order latent construct. However, for this type of formative model where the first-order dimensions have reflective indicators and the first-order dimensions act as formative indicators of the second-order construct, Jarvis et al. (2003) suggest that the influence of the component dimensions on the multidimensional composite construct must be non-contingent, that is, the dimensions must be independent of each other. However, in the case of Efficiency, System Availability, and Fulfilment, whilst it could be argued that System Availability is non-contingent with Efficiency and Fulfilment, System Availability impacts on both Efficiency and Fulfilment. Empirically, the non-contingency of the three dimensions is also hard to sustain given the relatively high correlations (see Table 3) between these dimensions and the high second-order factor loadings. As Fassnacht and Koese (2006) point out, the PZM study reports consistently high correlations in their two samples ranging from .62 to .78. According to Jarvis et al. (2003) high correlations between sub-dimensions pose serious measurement problems when they are modelled as formative indicators. This, therefore, suggests that both E-S-QUAL dimensions and their indicators are better modelled as reflective indicators of an overall e-service quality construct.

Following the above line of argument, a second-order measurement model was estimated for E-S-QUAL with scale items specified as reflective indicators of their respective E-S-QUAL dimensions and the E-S-QUAL dimensions specified as reflective indicators of a higher-order overall e-service quality construct. Although the second-order factor loadings were high (Efficiency had a factor loading of .94, System Availability .91, Fulfilment .70, and Privacy .65), the initial estimated model resulted in a poor level of fit: The observed χ^2 for this model is 922.60 and the χ^2/df ratio at 4.99 exceeded 3, as recommended by Bagozzi and Yi (1988). The GFI and adjusted GFI (AGFI) at .84 and .80 were much lower than the recommended level of .90. In an effort to address the problems, the modification indices (MI) were examined. First, two of the Efficiency items (EFF5: the website 'loads its pages fast'; and item EFF7: the website 'enables me to get on to it quickly') had errors that were correlated with the residual of the System Availability dimension, and as correlations with residuals are substantively uninterpretable, the model was re-estimated with the error covariance of items 5 and 7 of Efficiency specified as free parameters. Examination of the MI also showed that there was evidence of covariance between the residuals of Fulfilment and Privacy dimensions, and the regression weights of these dimensions also showed evidence of cross-loading. Such mis-specification means that the Fulfilment items could measure Privacy or vice versa. In Wolfinbarger and Gilly's (2003) study, privacy is identified as not being a significant factor in predicting e-service quality. This may also be the case in this study due to the high correlation between Fulfilment and Privacy. Also, Efficiency, System Availability, and Fulfilment could be regarded as *core* elements of the online (grocery) service, whilst Privacy is an augmented part of the service. Therefore, the model was re-estimated with Privacy specified as a free parameter.

The re-specified full measurement model yielded an overall chi square value of 373.01 with 101 degrees of freedom and $\chi^2/\text{df} = 3.69$. The GFI and AGFI were at .91 and .88. Although the value of AGFI was still below the recommended level of .90, other GFI appeared to be adequate (IFI, .96; CFI, .96; TLI, .95; and RMSEA, .08). All parameter estimates were statistically significant this time. The substantial

improvement in the model fit between the initial four-factor model and the final three-factor model suggests that Efficiency, System Availability, and Fulfilment are most appropriate for estimating E-S-QUAL in this study. All the model loadings were statistically significant, and the second-order loadings of the three dimensions are high: Efficiency had a regression weight of .95, System Availability .88, and Fulfilment .67. Chin (1998) suggests that for a second-order construct, a high proportion of the second-order paths should be $\geq$.70. In this study, two out of the three paths are >.7 (namely Efficiency and System Availability) and hence meet Chin's criteria. These findings together suggest the appropriateness of the second-order formulation of E-S-QUAL for assessing Internet retail service quality. To test for discriminant validity, average variance extracted (AVE) and shared variance were evaluated using the method advocated by Hair, Black, Babin, Anderson, & Tatham (2006). The AVEs ranged from .6 (Fulfilment) to .76 (Efficiency) and the shared variances (squared correlations) ranged from .34 to .70. Each AVE exceeded its respective shared variance between the factors, thereby satisfying the criteria for discriminant validity (Chin, 1998; Hair et al., 2006).

In order to verify the three-factor model of E-S-QUAL, the contribution of each of the E-S-QUAL factors was examined by regressing them on overall e-service quality. The summed scores of each of the factors served as independent variables, and overall Internet retail service quality was measured by asking the respondents to rate the performance of their website using the item 'The performance of this website meets my expectations'. Table 4 shows that Efficiency, System Availability, and Fulfilment were significant predictors of overall e-service quality but Privacy was not. Fulfilment had the strongest effect followed by System Availability and then Efficiency. This result is more consistent with the Wolfinbarger and Gilly's (2003) study than with the PZM study. In the PZM study, all four factors have a significant impact on quality, with Fulfilment having the strongest impact. In the Wolfinbarger and Gilly study, Fulfilment and website design (similar to the Efficiency and System Availability dimensions in E-S-QUAL) had a significant impact on quality but security/privacy did not. Notably, in all three studies, Fulfilment has the strongest impact on quality.

Nomological validity

The nomological validity of the resulting second-order model was tested by modelling E-S-QUAL as an antecedent of customer satisfaction and customer satisfaction as an antecedent of customer loyalty. This is based on strong empirical evidence in the offline context (see, e.g., Carrillat, Jaramillo, & Mulki, 2009) and emerging evidence

Table 4 Regression analysis of overall Internet retail service quality on E-S-QUAL dimensions.

Independent variables	Standardised coefficients			Collinearity statistics	
	β	*t*-values	Sig.	Tolerance	VIF
Privacy	.029	.578	.564	.512	1.955
Efficiency	.130	1.994	.047	.296	3.377
System availability	.195	3.155	.002	.330	3.028
Fulfilment	.371	6.967	.000	.446	2.241

Adjusted $R^2 = .393$ (sig. $p < .000$). Dependent variable: overall Internet retail service quality.

of the importance of e-service quality as an antecedent of customer satisfaction and customer loyalty in online retailing (see, e.g., Anderson & Srinivasan, 2003; Gefen, 2002; Shankar, Smith, & Rangaswamy, 2003). The three constructs were treated as latent constructs in a structural model. E-S-QUAL was modelled as a second-order latent construct with three dimensions. As previously mentioned, customer satisfaction was measured using the measure of online satisfaction employed by Jones and Suh (2000), and loyalty was measured using Zeithaml et al.'s (1996) five-item scale. The model showed a very good fit. The observed chi square for this model is 740.01, and the χ^2/df ratio is 2.99. The CFI at .952, NFI at .90, RFI at .92, and TLI at .95 are all above the recommended level of .90. RMSEA at .06 also suggests evidence of good fit. Furthermore, the standardised regression weight between E-S-QUAL and customer satisfaction was .56 and that between customer satisfaction and customer loyalty was .69. This provides strong support for the nomological validity of the three-dimensional, second-order formulation of E-S-QUAL. We tested the model further by adding an additional path from E-S-QUAL directly to loyalty. However, the path coefficient was relatively small (.084) and proved to be statistically insignificant ($p < .08$). This suggests that e-satisfaction acts as a mediator between e-service quality and e-loyalty.

Discussion and conclusions

E-S-QUAL provides a useful starting point for assessing Internet retail service quality. However, this study has shown that E-S-QUAL needs further refining. The results show that there are potential discriminant validity problems with the Efficiency and System Availability dimensions. The high correlation between the dimensions suggests that they are too closely correlated at the measurement level and therefore need more refinement to improve their discriminant validity.

Furthermore, whilst there is evidence that E-S-QUAL is a second-order construct, the analysis has shown potential problems of high correlation between the Privacy and Fulfilment dimensions. However, this does not mean that privacy is not important in predicting Internet retail service quality, given the moderately high correlations of Privacy with the other E-S-QUAL dimensions in both the PZM and this study. What this suggests, as Wolfinbarger and Gilly (2003) point out, is that Internet shoppers initially make inferences about privacy from other quality cues. In addition, for retailers that have both online and offline stores, Internet shoppers may be able to make inferences regarding privacy from the general reputation of the firm, as well as experience with loyalty schemes. It could also be argued that Efficiency, System Availability, and Fulfilment are the core dimensions of e-service quality and that assurances on Privacy are an order qualifying criteria for e-service patronage. That is, Privacy is something that the customers expect as a given before they buy from an Internet retailer. This is because privacy policies imply that an Internet retailer is trustworthy. Without privacy assurances, shoppers are less likely to complete the transaction (Elliott & Speck, 2005).

Interestingly, Fulfilment is the most important dimension of perceived e-service quality in both the PZM and this study. This underlines the importance of Fulfilment in the overall service quality outcome for Internet grocery shopping. The significantly lower regression coefficients of System Availability and Efficiency in this study suggest

that, overall, customers are relatively less happy with these dimensions, as in the PZM study, with Efficiency rated virtually equal with Fulfilment.

Managerial implications

The paper provides strong evidence that three-dimensional second-order version of E-S-QUAL (consisting of Efficiency, Fulfilment, and System Availability) displays robust psychometric properties and is a reliable measure of e-service quality. Whilst all three factors are critically important, for grocery e-tailers the study suggests that Fulfilment is the most important components of e-service quality because of the replenishment nature of grocery shopping. The relatively lower regression coefficients of Efficiency and Systems Availability suggest that e-tailers should be putting more of their efforts into website-related factors in the E-S-QUAL model. However, Privacy, whilst not a core dimension of e-service quality, is an important hygiene factor that customers expect as an essential prerequisite before purchasing from a website. Hence, e-tailers need to reassure customers continually of the privacy and security of their websites through appropriate website cues and other communication strategies.

Limitations

The findings of this study are limited to the e-grocery market, and therefore E-S-QUAL's psychometric properties need to be tested and validated further in other retail contexts in order to arrive at more comprehensive evaluation of the validity of E-S-QUAL and its dimensions. In particular, because grocery shopping is a replenishment, low-involvement, goal-directed activity, E-S-QUAL needs to be tested in more high-involvement, less goal-oriented shopping contexts such as fashion clothes shopping, for instance. In the high-involvement context, for instance, it is likely that non-service quality features (such as hedonic elements) of the website may be more important than in the low-involvement shopping context. This suggests that shopping orientation may act as a moderator of service quality perception, and this is therefore worth investigating. The study is also limited to just one country; a useful extension would be to expand the research into a number of other countries with different competitive, consumer, and business environments.

References

Anderson, R.E., & Srinivasan, S.S. (2003). E-satisfaction and e-loyalty: A contingency framework. *Psychology and Marketing*, *20*(2), 123–138.

Bagozzi, R.P., & Heatherton, T.F. (1994). A general approach to representing multifaceted personality constructs: Application to state self-esteem. *Structural Equation Modeling*, *1*(1), 35–67.

Bagozzi, R.P., & Yi, Y. (1988). On the evaluation of structural equation model. *Journal of Academy of Marketing Science*, *16*(1), 74–94.

Barnes, S.J., & Vidgen, R.T. (2002). An integrative approach to the assessment of e-commerce quality. *Journal of Electronic Commerce Research*, *3*(3), 114–127.

Boshoff, C. (2006). A psychometric assessment of E-S-QUAL: A scale to measure electronic service quality. *Journal of Electronic Commerce Research*, *8*(1), 101–114.

Brady, M.K., & Cronin, J.J., Jr. (2001). Some new thoughts on conceptualizing perceived service quality: A hierarchical approach. *Journal of Marketing*, *65*, 34–49.
Carrillat, F., Jaramillo, F., & Mulki, J. (2009). Examining the impact of service quality: A meta-analysis of empirical evidence. *Journal of Marketing Theory and Practice*, *17*(2), 95–110.
Chin, W.W. (1998). Issues and opinion on structural equation modeling. *MIS Quarterly*, *22*(1), 7–16.
Collier, J.E., & Bienstock, C.C. (2006). Measuring service quality in e-retailing. *Journal of Service Research*, *8*(3), 260–275.
Cristobal, E., Flavián, C., & Guinalíu, M. (2007). Perceived e-service quality (PeSQ); Measurement validation and effects on consumer satisfaction and web site loyalty. *Managing Service Quality*, *17*(3), 317–340.
Dabholkar, P.A., Thorpe, D.I., & Rentz, J.O. (1996). A measure of service quality for retail stores: Scale development and validation. *Journal of the Academy of Marketing Science*, *24*, 3–16.
Dagger, T.S., Sweeney, J.C., & Johnson, L.W. (2007). A hierarchical model of health service quality: Scale development and investigation of an integrated model. *Journal of Service Research*, *10*(2), 123–142.
Elliott, M.T., & Speck, P.S. (2005). Factors that affect attitude toward a retail web site. *Journal of Marketing Theory and Practice*, *13*(1), 40–51.
Fassnacht, M., & Koese, I. (2006). Quality of electronic services: Conceptualizing and testing a hierarchical model. *Journal of Service Research*, *9*(1), 19–37.
Garver, M.S., & Mentzer, J.T. (1999). Logistics research methods: Employing structural equation modeling to test for construct validity. *Journal of Business Logistics*, *20*(1), 33–57.
Gefen, D. (2002). Customer loyalty in e-commerce. *Journal of the Association for Information Systems*, *3*(1), 27–51.
Gerbing, D.W., & Anderson, J.C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, *25*(2), 186–192.
Grönroos, C. (1984). A service quality model and its marketing implications. *European Journal of Marketing*, *18*(4), 36–44.
Hair, J.F., Black, B., Babin, B., Anderson, R.E., & Tatham, R.L. (2006). *Multivariate data analysis* (6th ed.). London: Prentice-Hall.
Hubbard, R., Vetter, D.E., & Little, E.L. (1998). Replication in strategic management: Scientific testing for validity, generalizability, and usefulness. *Strategic Management Journal*, *9*(3), 243–254.
Janda, S., Trocchia, P.J., & Gwinner, K.P. (2002). Consumer perception of Internet retail service quality. *International Journal of Service Industry Management*, *13*(5), 412–431.
Jarvis, C.B., Mackenzie, S.B., & Podsakoff, P.M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*, 199–218.
Jones, M A., & Suh, J. (2000). Transaction-specific satisfaction and overall satisfaction: An empirical analysis. *The Journal of Services Marketing*, *14*(2), 147–159.
Loiacono, E.T., Watson, R.T., & Goodhue, D.L. (2002). WebQual: A measure of web site quality. In K.R. Evans & L.K. Scheer (Eds.), *Marketing Theory and Applications, 2002 Winter Educators' Conference, Proceedings* pp. 432–438. Chicago: American Marketing Association.
Nunnally, J.C. (1978). *Psychometric methods* (2nd ed.). New York: McGraw-Hill.
Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory*. New York: McGraw-Hill.
Oliver, R.L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, *17*(4), 460–469.
Parasuraman, A., Zeithaml, V.A., & Berry, L.L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, *64*(1), 12–40.

Parasuraman, A., Zeithaml, V.A., & Malhotra, A. (2005). E-S-QUAL: A multiple-item scale for assessing electronic service quality. *Journal of Service Research*, *7*(3), 213–233.
Rust, R., & Oliver, R.L. (1994). Service quality-insights and managerial implications from the frontier. In R.T. Rust & R.L. Oliver (Eds.), *Service quality: New directions in theory and practice* pp. 1–19. Thousand Oaks, CA: Sage.
Santos, J. (2003). E-service quality: A model of virtual service quality dimensions. *Managing Service Quality*, *13*(3), 233–246.
Shankar, V., Smith, A.K., & Rangaswamy, A. (2003). Customer satisfaction and loyalty in online and offline environments. *International Journal of Research in Marketing*, *20*(2), 153–175.
Spreng, R.A., & Mackoy, R.D. (1996). An empirical examination of a model of perceived service quality and satisfaction. *Journal of Retailing*, *72*(2), 201–214.
Steenkamp, J.-B.E.M., & van Trijp, H.C.M. (1991). The use of LISREL in validating marketing constructs. *International Journal of Research in Marketing*, *8*(4), 283–299.
Wolfinbarger, M., & Gilly, M.C. (2003). eTailQ: Dimensionalizing, measuring and predicting etail quality. *Journal of Retailing*, *79*(3), 183–198.
Yang, Z. (2001). Consumer perceptions of service quality in Internet-based electronic commerce. *Proceedings of the 30th EMAC Conference* pp. 8–11. Bergen, Norway: The Norwegian School of Economics & Business Administration.
Yoo, B., & Donthu, N. (2001). Developing a scale to measure the perceived quality of an Internet shopping site (SITEQUAL). *Quarterly Journal of Electronic Commerce*, *2*(1), 31–45.
Zeithaml, V.A., Berry, L.L., & Parasuraman, A. (1996). The behavioural consequences of service quality. *Journal of Marketing*, *60*(2), 31–47.
Zeithaml, V.A., Parasuraman, A., & Malhotra, A. (2000). *A conceptual framework for understanding e-service quality: Implications for future research and managerial practice* (Working Paper, Report No. 00–115). Cambridge, MA: Marketing Science Institute.
Zeithaml, V.A., Parasuraman, A., & Malhotra, A. (2002). Service quality delivery through web sites: A critical review of extant knowledge. *Journal of Academy of Marketing Science*, *30*(4), 362–375.

About the authors

Mohammed Rafiq is a reader in retailing and marketing at the Business School, Loughborough University. His current research interests include e-service quality, relationship quality, e-loyalty, internal marketing, market orientation, innovation, and new product development. His research has been published in the *Journal of Business Logistics*, *European Journal of Marketing*, *Journal of Services Marketing*, *European Journal of Innovation Management*, *European Journal of Information Systems*, *International Review of Retail, Distribution and Consumer Research*, among others. He served as co-editor of the *European Journal of Innovation Management* between 2006 and 2010. He has also co-edited a special issue of the *European Journal of Marketing* on the subject of internal marketing, and has chaired at various prestigious international conferences. Dr Rafiq has co-authored several books, including *Internal Marketing* and *Principles of Retail Management*.

Corresponding author: Mohammed Rafiq, Reader in Retailing and Marketing, Business School, Loughborough University, Leicestershire, LE11 3TU, UK.

T +44 (0) 1509 223397
E m.rafiq@lboro.ac.uk

Xiaoming Lu is currently a lecturer in marketing, travel, and tourism at Newcastle Business School, Northumbria University, UK. She obtained her doctorate from the Business School,

Loughborough University. Her current research interests are in the areas of the impact of technology on marketing communication, relationship quality, e-service quality, and e-loyalty.

T +44 (0) 191 227 3322
E lucy.lu@northumbria.ac.uk

Heather Fulford is currently a reader in entrepreneurship and academic director in the Centre for Entrepreneurship at Aberdeen Business School, Robert Gordon University. Additionally, she is a visiting fellow at the Business School, Loughborough University. Her main research interests are in the areas of IT in small businesses; electronic commerce adoption and diffusion in SMEs; innovation diffusion; language and translation technologies; website design and development; and information security management. Some of her recent research has been published in *Journal of Consumer Marketing*, *International Journal of Information Management*, *International Journal of Entrepreneurship Innovation*, *Journal of Enterprising Culture*, *International Journal of Retail and Distribution Management*, *Information Management and Computer Security*, *Information Resources Management Journal*, and *Computers and Security*. She is a member of the editorial advisory board of the *European Journal of Innovation Management*.

T +44 (0)1224 263869
E h.fulford@rgu.ac.uk